

# Control of multidimensional systems on complex network.

Giulia Cencetti,<sup>1,2,3</sup> Franco Bagnoli,<sup>2,3</sup> Giorgio Battistelli,<sup>1</sup> Luigi Chisci,<sup>1</sup> and Duccio Fanelli<sup>2,3</sup>

<sup>1</sup>*Dipartimento di Ingegneria dell'Informazione, Università di Firenze, Via S. Marta 3, 50139 Florence, Italy*

<sup>2</sup>*Dipartimento di Fisica e Astronomia and CSDC,*

*Università degli Studi di Firenze, via G. Sansone 1, 50019 Sesto Fiorentino, Italia*

<sup>3</sup>*INFN Sezione di Firenze, via G. Sansone 1, 50019 Sesto Fiorentino, Italia*

Multidimensional systems coupled via complex networks are widespread in nature and thus frequently invoked for a large plethora of interesting applications. From ecology to physics, individual entities in mutual interactions are grouped in families, homogeneous in kind. These latter interact selectively, through a sequence of self-consistently regulated steps, whose deeply rooted architecture is stored in the assigned matrix of connections. The asymptotic equilibrium eventually attained by the system, and its associated stability, can be assessed by employing standard nonlinear dynamics tools. For many practical applications, it is however important to externally drive the system towards a desired equilibrium, which is resilient, hence stable, to external perturbations. To this end we here consider a system made up of  $N$  interacting populations which evolve according to general rate equations, bearing attributes of universality. One species is added to the pool of interacting families and used as a dynamical controller to induce novel stable equilibria. Use can be made of the root locus method to shape the needed control, in terms of intrinsic reactivity and adopted protocol of injection. The proposed method is tested on both synthetic and real data, thus enabling to demonstrate its robustness and versatility.

PACS numbers: 89.75.Hc 89.75.Kd 89.75.Fb

Investigating the interlinked dynamics of an ensemble composed of units organized in homologous families, constitutes a universal challenge in science, of broad and cross-disciplinary breath [1–3]. Each population is customarily identified in terms of its continuous density. This latter evolves in time, as dictated by specific self-reaction stimuli, that generally bear nonlinear contributions. In a complex and dynamical environment, species experience a large plethora of mutual interactions, declinated via different modalities, notably pairwise exchanges. Cooperative and competitive interference are simultaneously at play, and shape the ultimate fate of the system as a whole [4]. These fundamental ingredients, flexibly combined and properly integrated, are at the roots of any plausible mathematical model targeted to community interactions [5], from ecology [6] to neuroscience [7, 8], passing from genetic and human health [9], through a full load of man-made technological applications [10]. Irrespectively of the specific realm of investigation, each population can be abstractly assigned to a given node of a virtual graph. Directed or undirected edges among nodes exemplify the topological structure of the existing network of interactions [11, 12]. The intricate web of inter-species connections, key information to anticipate the expected dynamics of the system, is therefore encoded in the associated adjacency matrix [13, 14].

In many cases of interest, it is important to drive the system towards a desired equilibrium, that is stable, and thus resilient, to external perturbations [15–19]. For example, hostile pathogens could be forced to go extinct: the stability of the attained equilibrium would efficaciously shield from subsequent harmful invasion and outbreaks. Alternatively, it could prove vital to robustly enhance the expression of species identified as beneficial for the system at hand. Building on these premises, we here develop and test a general control strategy [20–24] targeted to multidimensional systems consisting of a large number of components that interact through a complex network. By inserting one additional species, the controller, which configures as a further node of the collection, we will be able to manipulate the asymptotic dynamics of the system, in terms of existence and stability of the allowed fixed points.

To set the reference frame we will hereafter consider a system consisting of  $N$  species (nodes) whose activities  $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$  obey the coupled nonlinear equations [6, 15, 16]:

$$\dot{x}_i = f_i(x_i) + \sum_j A_{ij} g_i(x_i, x_j) \quad i = 1, \dots, N. \quad (1)$$

The first term on the right-hand side specifies the self-dynamics of species  $i$  while the second term stems from the interactions of species  $i$  with the other species. The nonlinear functions  $f_i(x_i)$  and  $g_i(x_i, x_j)$  encode the dynamical laws that govern the system's components, while the weighted connectivity matrix  $\mathbf{A}$  captures the interactions between nodes. The elements  $A_{ij}$  can be positive or negative, depending on the specific nature of the interaction, i.e. cooperative or competitive. Notice that system (1) is assumed in [15] as a reference model to analyze resilience patterns in complex networks. Differently from [15],  $A_{ij}$  can here take positive and negative values (see also [25]).

In ecological applications, the number of nodes reflects the biodiversity of the scrutinized sample [6, 16, 26]. Distinct trophic layers materialize as coherent blocks in the adjacency matrix, whose entries modulate the strength of mutual interactions [4]. These are often epitomized by a quadratic response function  $g_i(x_i, x_j)$  [16]. Each species is then subjected to a reaction drive  $f_i(x_i)$ , usually a logistic growth with a prescribed carrying capacity [15]. Animals displaying competitive predator-prey interactions or, alternatively, subjected to a symbiotic dependence, such as in plant-pollinator relationships, are among the systems that fall within the aforementioned scenario [4]. Furthermore, the complex community of micro-organisms that live in the digestive tracts of humans and other animals, including insects, can be rooted on similar descriptive grounds [6]. For genetic regulatory networks, the dynamical variables  $x_i$  represent the level of activity of a gene or the concentration of the associated proteins [27]. Species specific reaction terms  $f_i(x_i)$  account for e.g. degradation or dimerization. The pattern of activation could be effectively modeled by sigmoidal Hill-like functions [1], as follows the classical Michaelis-Menten scheme [28], which incorporates the known map of gene interactions. On a more general perspective, understanding the emerging dynamics in social communities [29], grasping the essence of the learning organization in the brain [30], and implementing efficient protocols for robot navigation in networked swarms [31] are among the very many applications that can be traced back to one of the variants of equations (1), with a suitable choice of the nonlinear functions  $f_i(x_i)$  and  $g_i(x_i, x_j)$ .

### Adding a species to enforce stable equilibria in a multidimensional system

Starting from the above illustrated setting, we will here discuss a suitable control scheme to drive system (1) towards a desired equilibrium  $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_N^*)^T$ , which is linearly stable to externally imposed perturbations. To reach this goal we shall introduce one additional species, the  $(N + 1)$ -th component of the collection, suitably designed to yield the sought effect. To set the notation, we indicate by  $u$  the component (e.g., concentration, activation level) assigned to the controller and write:

$$\begin{cases} \dot{x}_i = f_i(x_i) + \sum_j A_{ij} g_i(x_i, x_j) + \alpha_i h_i(x_i, u) \\ \dot{u} = -(u - u^*) - \rho \sum_j \beta_j (x_j - x_j^*) \end{cases} \quad (2)$$

The controller  $u$  can exert a direct influence on every component  $x_i$ , as specified by newly added terms  $\alpha_i h_i(x_i, u)$  that modify the original system (1).  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$  is a vector of  $N$  constant parameters, to be self-consistently adjusted following the scheme depicted below.  $h_i(x_i, u)$  is a generic, in principle nonlinear, function of the components  $x_i$  and  $u$  that reflects the modality of interactions between the controller and the existing species. The equation for the dynamical evolution of the controller  $u$  displays two distinct contributions. The first represents a self-reaction term, assumed to be linear just for ease of presentation. The nonlinear self-dynamics of the controller  $u$  can be readily considered, with no further technical complication. The rate of change of  $u$  is assumed to be contextually driven by a global forcing that senses the relative distance of  $x_i$  from its deputed equilibrium  $x_i^*$ . The parameters  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_N)^T$  and  $\rho$  will prove central in enforcing the stabilization of the prescribed fixed point. A few comments are mandatory to fully appreciate the generality of the proposed framework, beyond the specific choices made for purely demonstrative purposes. Let us begin by remarking that the controller  $u$  can represent an artificially engineered component or, equivalently, belong to an extended pool of interacting populations. In the scheme here imagined, it is assumed that the values of  $u$  and  $x_i \forall i$ , are accessible to direct measurements at any time and that this information can be processed to set the controller dynamics. This is largely reasonable for experiments that run under protected conditions like, e.g., the study of microbial dynamics in laboratory reactors, but certainly less realistic for applications that aim at in vivo multidimensional systems, think for instance to genetic regulatory circuits. The dynamical equation for  $u$  can, however, be amended to a large extent and with a great deal of flexibility, depending on the target application and the structural specificity of the employed controller, while still allowing for an analogous methodological treatment<sup>1</sup>. The dynamics of the original, unsolicited, components and the functional form that specifies the controller feedback bear unequivocal universality traits [15].

The global fixed point  $(\mathbf{x}^*, u^*)$  of the controlled system (2) should match the following constraints

---

<sup>1</sup> As a matter of fact, we can equivalently assume a generalized equation for the controller of the type  $\dot{u} = f_u(u) - \rho g_u(\mathbf{x}, u, \boldsymbol{\beta}) + b$  where  $f_u(u^*) = 0$  and  $b = \rho g_u(\mathbf{x}^*, u^*, \boldsymbol{\beta})$ .

$$f_i(x_i^*) + \sum_j A_{ij} g_j(x_i^*, x_j^*) + \alpha_i h_i(x_i^*, u^*) = 0 \quad i = 1, \dots, N \quad (3)$$

which, provided the  $x_i^*$  and  $u^*$  are assigned, ultimately set the values of the parameters  $\alpha_i$ . Conversely, as we shall illustrate in the following, one could assume the parameters  $\alpha_i$  as a priori known and infer via equations (3) the fixed point(s) to be eventually stabilized. The next step in the analysis aims at ensuring the stability of the selected fixed point. This will be achieved by acting on the residual free parameters  $\beta$  and  $\rho$ . As routinely done, we perturb the equilibrium solution as  $x_i = x_i^* + v_i$ ,  $u = u^* + w$  and Taylor expand equations (2) assuming the imposed disturbances  $\eta = (v, w)$  small in magnitude. At the linear order of approximation one obtains:

$$\dot{\eta} = \begin{pmatrix} \mathbf{G} & \mathbf{q} \\ -\rho \beta^T & -1 \end{pmatrix} \eta \equiv \mathbf{J} \eta \quad (4)$$

where  $\mathbf{q}$  is a  $N$ -dimensional column vector of components  $q_i = \alpha_i \frac{\partial h_i}{\partial u}(x_i^*, u^*)$ . The  $N \times N$  matrix  $\mathbf{G}$  is defined as:

$$\begin{aligned} G_{ii} &= \frac{\partial f_i}{\partial x_i}(x_i^*) + \sum_k A_{ik} \frac{\partial g_k}{\partial x_i}(x_i^*, x_k^*) + \alpha_i \frac{\partial h}{\partial x_i}(x_i^*, u^*) \\ G_{ij} &= A_{ij} \frac{\partial g_j}{\partial x_j}(x_i^*, x_j^*). \end{aligned}$$

The fixed point  $(\mathbf{x}^*, u^*)$  is linearly stable if all eigenvalues of the Jacobian matrix  $\mathbf{J}$  have negative real parts. The associated characteristic polynomial  $P(\lambda) = \det(\mathbf{J} - \lambda \mathbf{I})$  can be cast in the equivalent, affine in the  $\rho$ -parameter, form:

$$P(\lambda) = (-1)^{N+1} (1 + \lambda) \det(\mathbf{G} - \lambda \mathbf{I}) - \rho \sum_{i,j=1}^N \beta_j [\text{adj}(\mathbf{G} - \lambda \mathbf{I})]_{ji} q_i \equiv \mathcal{D}(\lambda) + \rho \mathcal{N}(\lambda)$$

that is reminiscent of the celebrated root locus method [32]. Here,  $[\text{adj}(\mathbf{Z})]_{ji} = (-1)^{i+j} \det[(\mathbf{Z})_{(i,j)}]$  denotes the adjugate of matrix  $\mathbf{Z}$ ,  $(\mathbf{Z})_{(i,j)}$  being the minor of  $\mathbf{Z}$  obtained by removing the  $i$ -th row and the  $j$ -th column. The polynomials  $\mathcal{D}(\lambda) = -\prod_{k=1}^{N+1} (\lambda - p_k)$ , and  $\mathcal{N}(\lambda) = \prod_{k=1}^{N-1} (\lambda - z_k)$  have respectively degrees  $N + 1$  and  $N - 1$ . With a slight abuse of language we will refer to as *poles* the roots  $p_k$  of the polynomial  $\mathcal{D}(\lambda)$  and *zeros* the roots  $z_k$  of  $\mathcal{N}(\lambda)$ . Notice that for  $\rho = 0$  the eigenvalues of the Jacobian  $\mathbf{J}$  correspond to the  $N + 1$  poles  $p_k$ . These latter quantities are uniquely determined, once the fixed point  $(\mathbf{x}^*, u^*)$  has been assigned. In particular it cannot a priori be ensured that the real parts of all  $p_k$  are negative, as stability would require. In other words, when  $\rho = 0$ , we can enforce the desired fixed point into the system but cannot guarantee its stability. On the other hand, for  $\rho \rightarrow \pm\infty$ ,  $N - 1$  eigenvalues of  $\mathbf{J}$  tend to the zeros  $z_k$ , which depend self-consistently on the free parameters  $\beta$ . As we shall show hereafter, it is in principle possible to assign the  $\beta_i$  to force the real parts of all  $z_k$  to be negative. The two remaining eigenvalues of matrix  $\mathbf{J}$ , in the limit of large  $|\rho|$ , diverge to infinity in the complex plane. More precisely, they travel along opposite directions following a vertical (resp. horizontal) asymptote, if  $\rho$  is bound to the positive (resp. negative) semiaxis. To confer stability in the limiting case  $\rho \rightarrow \infty$  where  $N - 1$  eigenvalues of  $\mathbf{J}$  coincide with the roots of  $\mathcal{N}(\lambda)$ , it is therefore sufficient to (i) operate a supervised choice of  $\beta$  and (ii) impose the condition that yields a vertical asymptote ( $\rho \rightarrow +\infty$ ), while, at the same time, requiring that this latter intersects the negative side of the real axis. In this respect, it is important to remark that the intersection occurs in the point of abscissa  $\lambda_0 = \frac{1}{2} (\sum_{k=1}^{N+1} p_k - \sum_{k=1}^{N-1} z_k)$ . Hence, the idea is to interpolate between the two limiting cases  $\rho = 0$  and  $\rho \rightarrow \infty$  by determining the minimal value  $\rho_c$  of  $\rho$  beyond which the desired fixed point becomes stable. The existence of the threshold  $\rho_c$  that makes the imposed fixed point attractive for any  $\rho > \rho_c$  is obvious, being stability already assured in the limiting setting  $\rho \rightarrow +\infty$ <sup>2</sup>. For the sake of clarity we reiterate that this amounts to selecting  $\text{Re}(z_k) < 0$  for all  $k$  and further imposing  $\lambda_0 < 0$ , by properly assigning the free parameters  $\beta$ .

<sup>2</sup> In principle, more than one value of  $\rho_c$  can exist for which the eigenvalues cross the imaginary axis, making stable an unstable fixed point. The intersections are found imposing  $\lambda = i\omega$  in equation  $\mathcal{D}(\lambda) + \rho_c \mathcal{N}(\lambda) = 0$ , which yields a system of two equations, for respectively the real and imaginary parts. This system can then be solved for the two unknowns  $\rho_c$  and  $\omega$ .

In order to study the assignability of the zeros  $z_k$  by means of  $\beta$ , let us recall that for a generic square matrix  $\mathbf{Z}$ ,  $\text{adj}(\mathbf{Z} - \lambda \mathbf{I}) = -\sum_{m=0}^{N-1} \sum_{l=0}^{N-m-1} c_{l+m+2} \mathbf{Z}^m \lambda^l$  where  $c_k$  stands for the coefficients of the characteristic polynomial of  $\mathbf{Z}$ , namely  $\det(\mathbf{Z} - t \mathbf{I}) = \sum_{l=0}^N c_{l+1} t^l$ . The polynomial  $\mathcal{N}(\lambda)$  can be consequently written as:

$$\mathcal{N}(\lambda) = \sum_{m=0}^{N-1} \sum_{l=0}^{N-m-1} c_{l+m+2} [\beta^T \mathbf{G}^m \mathbf{q}] \lambda^l \equiv \sum_{n=0}^{N-1} d_{n+1} \lambda^n \quad (5)$$

It is hence straightforward to establish a direct relation between the parameters  $\beta$  and the vector of coefficients  $\mathbf{d}$ :

$$d_n = \sum_{k=0}^{N-n} c_{k+n+1} [\beta^T \mathbf{G}^k \mathbf{q}] \quad (6)$$

that can also be equivalently stated as:

$$\mathbf{d} = \mathbf{H} \beta \quad (7)$$

where  $\mathbf{H}$  is the matrix defined by:

$$H_{nm} = \sum_{k=0}^{N-n} c_{k+n+1} (\mathbf{G}^k \mathbf{q})_m. \quad (8)$$

The suited vector  $\beta$  is thus obtained<sup>3</sup> from (7), provided matrix  $\mathbf{H}$  is invertible. This latter request defines the condition of *controllability* for the scheme that we have implemented (see Supplementary Information, SI, for a discussion that aims at positioning this observation in the context of standard control theory [33]). Summing up, the devised strategy consists of the following steps. First, the fixed point is selected and the parameters  $\alpha$  frozen to their respective values as specified by Eq. (3). Then the complex roots  $z_k$  are chosen so that  $\text{Re}(z_k) < 0$  for all  $k$  while, at the same time, matching the condition that makes the vertical asymptote cross the horizontal axis with a negative intercept. As we will clarify when discussing the applications, the  $z_k$  can be chosen to coincide with the poles  $p_k$ , except for punctual modifications whenever  $\text{Re}(p_k) > 0$ . Notice however that  $z_k$  should be real or come in conjugate pairs, as the coefficients  $d_k$  are, by definition, real. Once the roots  $z_k$  have been fixed, one can readily compute the associated polynomial coefficients  $d_k$ , and hence proceed with the determination of  $\beta$  via (7), provided that the controllability condition holds. Finally, by selecting  $\rho > \rho_c > 0$  we obtain a linearly stable fixed point  $(\mathbf{x}^*, u^*)$  for the controlled dynamics (2).

### Testing the control method: from synthetic gene network to real microbiota dataset

As a first application of the above technique, we will study the dynamics of an artificial gene network [34–37]. In our example the network of connections is a regular tree with branching ratio  $r = 4$ . It is further assumed that the genetic activation between nodes  $i$  and  $j$  is described in terms of a Hill function, with cooperation coefficient equal to 2. In formulae,  $A_{ij} = 1$  and  $g_i(x_i, x_j) \equiv g(x_j) = x_j^2 / (1 + x_j^2)$ . Negative regulation loops are also accommodated for. These latter could, in principle, be modeled by assuming paired interactions of the type  $1 - g(x_j)$ , while still setting to one the relative entry of the connection matrix. As described in the SI, we can equivalently set  $A_{ij} = -1$ , while assuming interactions to be modulated by  $g(x_j)$  as indicated above. At the same time, the reaction part should be modified with an additional term,  $\eta_i$ , counting the number of negative loops that affects node  $i$ . More specifically,  $f(x_i) = -\gamma_i x_i + \eta_i$ , where the first term mimics constitutive degradation. In our tests, matrix  $\mathbf{A}$  contains an identical number of randomly assigned  $\pm 1$ . The parameters  $\gamma_i$  are random variables uniformly distributed over the interval  $[0, 1]$ . Working in this setting our aim is to control the equilibrium state of the system and thus shape the pattern of asymptotic activity. For this initial application we choose to operate with a simple linear control and set  $h_i(x_i, u) \equiv h(u) = u$ , for all  $i$ . In this case,  $u$  could e.g. represent the density of a suitable retroviral vector used to infect specific cell lines [38]. To provide

---

<sup>3</sup> For obvious consistency reasons  $\beta$  must have real entries. This follows naturally if one chooses the zeros  $z_k$  to be real or complex conjugate in pairs, which implies that the coefficients  $d_n$  of the polynomial  $\mathcal{N}(\lambda)$  (see (5)) are real. All other quantities involved are real by definition.

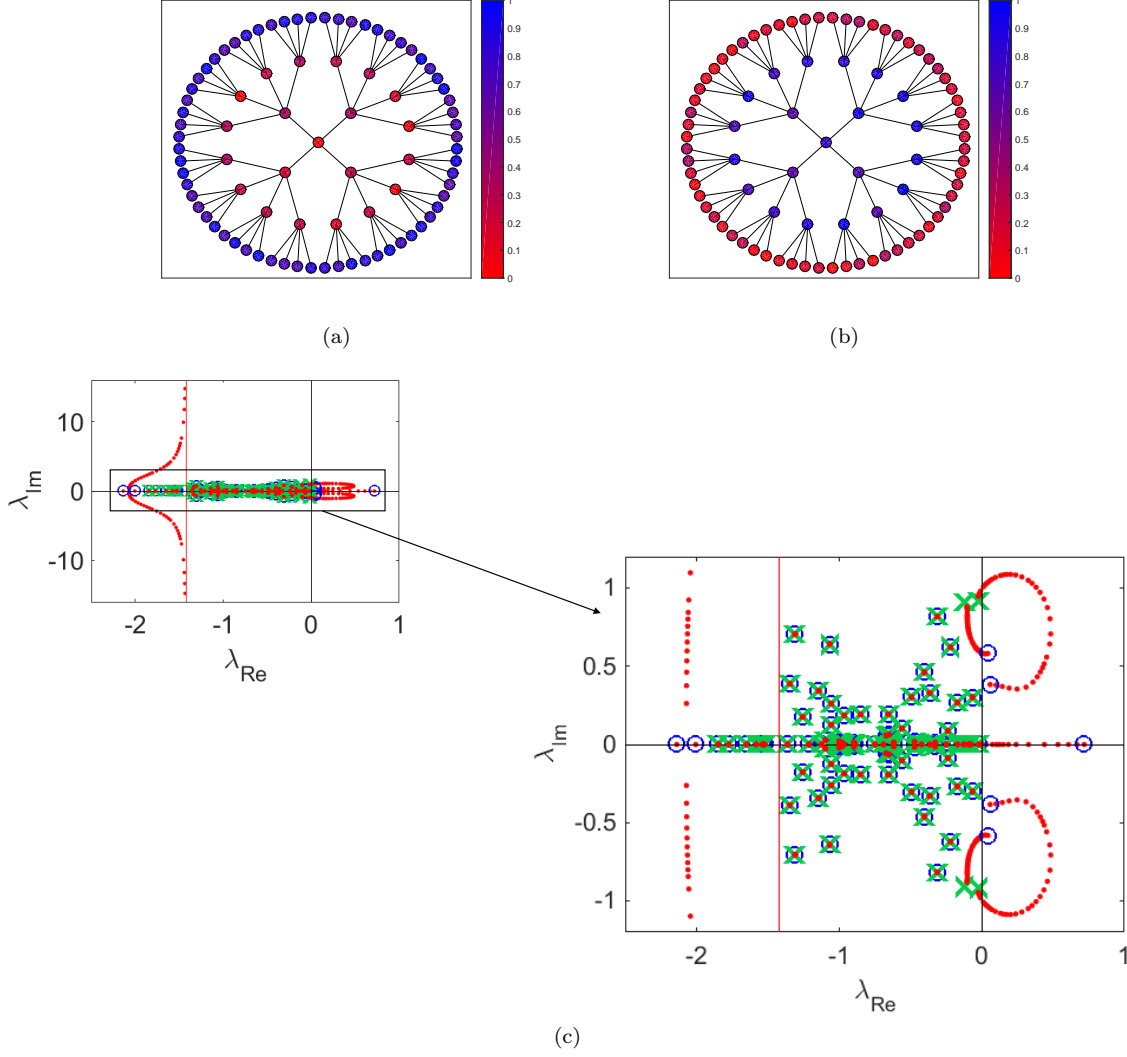


FIG. 1: Panel (a): the control is modulated so as to enhance the activity of the peripheral nodes of the tree, as compared to the inner ones. Panel (b): the control makes now the bulk nodes more active as compared to the peripheral ones. Panel (c): the root locus diagram relative to the situation displayed in panel (b) is plotted. Blue circles stand for the position of the complex eigenvalues when  $\rho = 0$ , while green crosses identify the eigenvalue obtained for  $\rho \rightarrow \infty$ . The vertical red line represents the asymptote that attracts two of the modified eigenvalues, when  $\rho \rightarrow \infty$ . The red dots show the computed spectrum, calculated when increasing  $\rho$ . In this case the matrix  $\mathbf{A}$  contains an identical number of  $\pm 1$  entries. These are randomly assigned and kept unchanged for all tests performed. The figure on the right is a zoom of the plot displayed on the left.

an immediate graphical illustration of the power of the method, we set to stabilize two distinct fixed points. In the first example, see Figure 1(a), the control is designated so as to enhance the degree of activity of the peripheral nodes of the tree. These latter are characterized by a similar value of the activity, apart for slight randomly superposed fluctuations. Similarly, the nodes that define the bulk of the tree display a shared degree (except for tiny stochastic modulation) of residual activity. In Figure 1(b), the dual pattern is instead obtained and stabilized: the peripheral nodes are now being silenced and the activity concerns the nodes that fall in the center of the tree. In Figure 1(c) the root locus diagram relative to the situation reported in Figure 1(b) is displayed. By properly tuning  $\rho$  above a critical threshold  $\rho_c$ , we can enforce the stability of the obtained fixed point. Two eigenvalues diverge to  $\pm\infty$  following a vertical asymptote in the complex plane. For each chosen fixed point that is being stabilized the zeros  $z_k$  can be selected so as to make the asymptote intercept the horizontal axis in the left-half of the plane.

As a second application of the proposed control strategy, we set to study the dynamics of the gut microbiota [6]. The intestinal microbiota is a microbial ecosystem of paramount importance to human health [39]. Efforts are currently

aimed at understanding the microbiota ability to resist to enteric pathogens and assess the response to antibiotics cure of intestinal infections. Recent advances in DNA sequencing and metagenomics make it possible to quantitatively characterize the networks of interactions that rule the dynamics of the microbiota ecosystem. This was for instance achieved in [40] by analyzing available data on mice [41] with an innovative approach which combines classical Lotka-Volterra model and regression techniques. Eleven species were identified and thoroughly analyzed in terms of self and mutual dynamics.

In the following we shall apply the method here developed to control the dynamics of the whole microbioma [40] or a limited sub-portion of it. In this specific case, the self-dynamics is assumed to be logistic, namely  $f_i(x_i) = x_i(r_i - s_i x_i)$ , while  $g(x_i, x_j) = x_i x_j$ . The constants  $r_i$  and  $s_i$  are provided in [40] and follow from direct measurements. The weighted matrix of connections  $\mathbf{A}$  presents both positive and negative entries, assigned according to [40]. Finally,  $h_i(x_i, u) = u x_i$ . The results of the analysis are organized under different headings that reflect the three distinct control strategies explored.

**Stabilizing an unstable fixed point by means of an external controller (Case A).** Consider the system of 11 species, as defined in [40] (see SI for a discussion on the bacterial species involved). For illustrative purposes, we will restrict the analysis to all sub-systems that combine 5 out of the 11 species analyzed in [40]. The fixed points for the obtained 5 species systems are calculated. Those displaying positive concentrations are then retained for subsequent analysis. The stability of each selected fixed point is established upon evaluation of the spectrum of the Jacobian of the reduced dynamics. In Figure 2(a) the histogram of  $(\lambda_{Re})_{max}$ , the largest real parts of the recorded eigenvalues, is plotted: several fixed points exist that correspond to unstable equilibria. Starting from this setting, we will introduce a suitably shaped controller, following the above discussed guidelines, in order to stabilize a slightly perturbed version of an originally unstable fixed point, see pie charts in Figure 2(a). Denote by  $\mathbf{x}^*$  the fixed point to be eventually stabilized and consequently assign the parameters  $\alpha_i$  so as to match Eqs. (3). The spectrum of the Jacobian matrix obtained for  $\rho = 0$  (blue circles in Figure 2) protrudes into the right half-plane. More specifically, one eigenvalue exhibits a positive real part, so flagging the instability that one aims to control. At variance, the crosses in Figure 2(b) stand for the roots  $z_k$  of  $\mathcal{N}(\lambda)$ , and fall in the left side of the complex plane. The vertical (red, in Figure 2(b)) line identifies the location of the two residual eigenvalues of the Jacobian matrix, when  $\rho \rightarrow \infty$ . By tuning the parameter  $\rho$ , one can continuously bridge the two above limiting settings, as graphically illustrated in Figure 2(b). When  $\rho > \rho_c \simeq 0.01$ , the eigenvalues populate the left half-hand plane and stability is, therefore, gained.

**Acting with one species of the pool to damp the concentration of the pathogens (Case B).** Select now a stable fixed point, mixture of five distinct species. One of them is *Clostridium difficile*, a species of Gram-positive spore-forming bacteria that may opportunistically dominate the gut flora, as an adverse effect of antibiotic therapy. As controller we shall here employ one of the other 6 species that compose the microbioma [42, 43]. The aim is to drive the system towards another equilibrium, stable to linear perturbations, which displays a decreased pathogen concentration. In this case the parameters  $\alpha$  are determined a priori, once the control species has been identified. Denote by  $\bar{\mathbf{A}}$  the reduced  $5 \times 5$  matrix that specifies all paired interactions between the pool of populations involved in the initial fixed point. The equilibrium solution that can be attained by the controlled system is determined as  $\mathbf{x}^* = -\bar{\mathbf{A}}^{-1}(\mathbf{r} + \alpha)$ , and clearly depends on the species used as controller. The only meaningful solutions are those displaying non negative components  $x_i^*$ . In the example depicted in Figure 2(c) only three solutions can be retained, namely the ones obtained by using uncl. Lachnospiraceae, uncl. Mollicutes and *Enterococcus* as respective control. In one of the inspected cases, the amount of *C. difficile* is found to reduce, when the control is turned on. The asymptotic concentration that is eventually attained is sensibly lower than the one initially displayed. The pie charts in Figure 2(c) represent, respectively, the initial fixed point and the final stationary equilibrium, as shaped by the control in the most beneficial case, i.e., when the concentration of *C. difficile* is seen to shrink. The root locus plot obtained for this specific case is reported in the SI. Importantly, the discussed scheme can be straightforwardly modified so as to account for a generic nonlinear self-reaction dynamics for the control species, e.g., a logistic growth, that could replace the linear Hookean-like term assumed in Eq. (2).

**Driving to extinction one species, the other being the target of the control (Case C).** As an additional example, we wish to modify a stable fixed point of the dynamics, by silencing one of the existing populations with an

indirect control. In other words we shall introduce and stabilize a novel fixed point, that displays a negligible residual concentration of the undesired species, by acting on the other species of the collection. This is for instance relevant when aiming at, e.g., eradicating a harmful infection that proves resistant to direct therapy. With this in mind, we consider a reduced ecosystem consisting of 6 species, selected among the 11 that define the microbiota. A stable fixed point exists (black diamonds in Figure 2(d), left panel) which displays a significant concentration of *C. difficile*, the pathogen species. Assign to this latter species the index 6. We now insert a controller which cannot directly interfere with *C. difficile*. This amounts, in turn, to setting to zero the corresponding component of vector  $\alpha$  ( $\alpha_6 = 0$ ). We then require the concentration of the *C. difficile* to be small, i.e.,  $x_6^* = \varepsilon \ll 1$ . This latter condition translates into a constraint that should be matched by the other 5 species, namely  $\sum_{j \neq 6} \bar{A}_{6j} x_j^* = (s_6 - \bar{A}_{66})\varepsilon - r_6$ . Given  $x_k^*$ , the components  $\alpha_k$ , with  $k \neq 6$ , are chosen so as to match the constraint  $\alpha_k = (-r_k + s_k x_k^* - \sum_{j \neq 6} \bar{A}_{kj} x_j^* - \bar{A}_{k6}\varepsilon)/u^*$ . A possible solution of the problem is reported in Figure 2(d): in the left panel (plus symbols) the components of the fixed point stabilized by the control are shown. As anticipated, the concentration of *C. difficile* is small. The right panel of Figure 2(d) shows the components of the vector  $\alpha$  that specify the characteristics of the introduced controller. Notice that  $\alpha_6 = 0$  so that the controller is not directly influencing the rate of production of *C. difficile*.

## Discussion

We would like to draw the attention on the interpretation of  $\alpha$ . As stated earlier,  $\alpha$  characterizes the strength of the coupling between the controller  $u$  and every single species of the system to be controlled. An alternative interpretation is however possible:  $u$  could represent a mixture of different species and the components of  $\alpha$  incorporate the relative abundance of the mixed compounds. In light of the above, also the previously discussed control schemes which apparently assumed dealing with an artificially designed control, could be realized via a proper mixture of existing microbiota species so as to achieve the coupling  $\alpha$  corresponding to the desired fixed point.

Notice also that the control scheme here developed could be in principle exploited to drive the system towards a stable fixed point of the unperturbed dynamics, starting from out-of-equilibrium initial conditions. To achieve this goal  $u^*$  needs to be set to zero, thus requiring that the controller is turned off at equilibrium. In this case,  $\alpha$  and  $\beta$  are not subjected to specific constraints, as the existence and stability of the desired equilibrium are a priori granted. Such parameters could hence be chosen so as to reflect the specificity of the target system. In the annexed SI we demonstrate this intriguing possibility.

Summing up, we have here proposed and tested a method to control the dynamics of multidimensional systems on a complex graph. The original system is made up of  $N$  interacting populations obeying a set of general equations, which bear attributes of universality. One additional species, here referred to as the controller, is inserted and made interact with the existing constellation of species. By tuning the strength of the couplings (or equivalently the composition of the inserted controller), we can drive the system towards a desired equilibrium. The stability of the achieved solution is enforced by adjusting the parameters that ultimately govern the rate of change of the controller. Methodologically, we make use of the root locus method which can be naturally invoked once the control problem is suitably formulated. The tests that we have performed, both synthetic and drawn from real life applications, demonstrate the versatility and robustness of the proposed scheme. This latter configures therefore as a viable and innovative tool to tackle a large plethora of inter-disciplinary systems, from life science to man-made applications, that should be stably driven towards a desired configuration. In this current implementation, and for purely pedagogical reasons, the control assumes that the state of the system is accessible to direct measurement. Relaxing this working hypothesis is a possibility that we shall explore in a future contribution.

## SUPPLEMENTARY INFORMATION

### Genetic network model

We shall here justify the model of genetic regulatory network analyzed in the main text. Consider first a small regulatory network consisting of one gene (whose activity is labelled  $x$ ) and one protein (associated to the continuous concentration  $y$ ). A positive regulation loop can be modeled as:

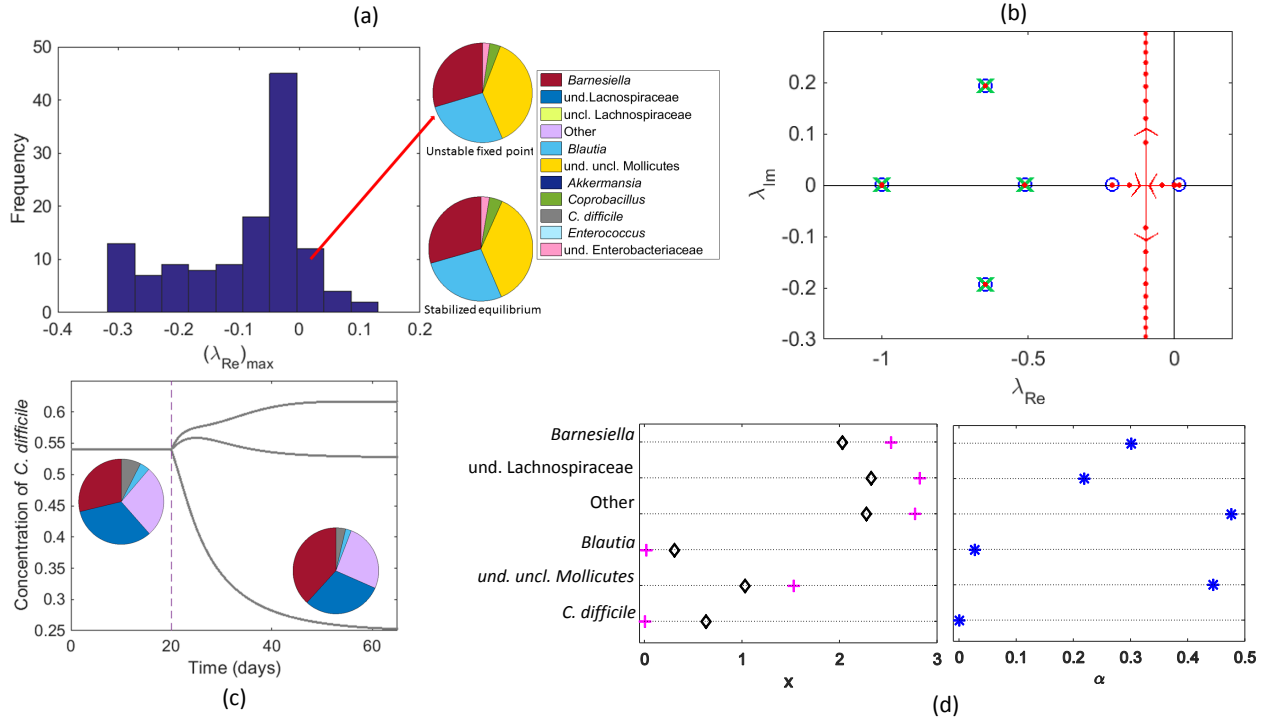


FIG. 2: Panel (a): a reduced 5 species subsystem of the microbiota is considered (case A) and all possible fixed points computed. Only those displaying non-negative concentrations are retained and their stability assessed. In the main figure, the histogram of  $(\lambda_{Re})_{max}$ , the largest real parts of the eigenvalues obtained after the linear stability analysis, is depicted. The two pie charts refer to the initially unstable fixed point (upper chart) and the stabilized equilibrium (lower chart). Panel (b): the root loci diagram relative to the case discussed in panel (a), is shown. Blue circles identify the position of the complex eigenvalues when  $\rho = 0$ , while green crosses stand for the eigenvalues obtained in the limit  $\rho \rightarrow \infty$ . The vertical red line is the asymptote that eventually attracts the two residual eigenvalues. The red dots show the computed spectrum, when progressively increasing  $\rho$ . Panel (c): the goal is here to reduce the concentration of the pathogen species, *C. difficile*, by employing as controller one of the species that compose the microbioma (case B). The concentration of *C. difficile* is monitored over time for three different control strategies, turning on the control at the same time ( $t = 20$  days). The insertion of the species of und. Lachnospiraceae provokes a substantial reduction (50 %) of the pathogen concentration, as also displayed by the enclosed pie charts (for interpreting the color-code refer to panel (a)). Panel (d): we now modify a stable fixed point, by driving to extinction one of the existing population, the pathogen *C. difficile* (here species 6), with an indirect control strategy (case C). The obtained concentrations are reported in the left graph (pluses) and confronted with the initial unperturbed solution (diamonds). As anticipated  $x_6^* \simeq 0$ . The components of  $\alpha$  are plotted in the right graph. Notice in particular that  $\alpha_6 = 0$ .

$$\begin{aligned}\dot{x} &= k_1 g(y) - \gamma_1 x \\ \dot{y} &= k_2 x - \gamma_2 y\end{aligned}$$

where:

$$g(y) = \frac{y^n}{K + y^n}. \quad (S9)$$

In the following, we will make the choice  $n = 2$  and  $K = 1$ . Similarly, a negative regulation loop can be modeled as:

$$\begin{aligned}\dot{x} &= k_1(1 - g(y)) - \gamma_1 x \\ \dot{y} &= k_2 x - \gamma_2 y\end{aligned}$$

As it should be, the concentration of proteins grows with the level of gene activity. As a first step approximation, and to avoid dealing with two distinct families of mutually interlinked constituents (genes and proteins), we can



replace  $y$  with  $x$  in the argument of  $g(\cdot)$ , via adiabatic elimination (apart from redefinition of the involved constants). Building on the above we model the extended regulatory network as:

$$\dot{x}_i = k_i \sum_j A_{ij} g(x_j) - \gamma_i x_i + k_i \eta_i \quad (\text{S10})$$

where  $x_i$  stands for the activity of gene  $i$ . For *positive* feedbacks between species  $i$  and  $j$ ,  $A_{ij} = 1$ , while, for *negative* loops  $A_{ij} = -1$ . The parameter  $\eta_i$  stands for the number of negative loops (number of negative entries of the  $i$ -th row of  $\mathbf{A}$ ) that are associated to node  $i$ . Finally, to keep the structure as simple as possible we set  $k_i = 1$  for all  $i$ . The matrix  $\mathbf{A}$  employed in the example reported in the main body has a simple structure and it has been chosen for purely illustrative purposes: it represents a regular tree network with branching ratio  $r = 4$ .

### Controlling the Microbiota network: the experimental data

The second application of the proposed control method deals with a model of gut microbiota. As explained in the main body of the paper, the model is well established and builds on a generalization of the celebrated Lotka-Volterra equations:

$$\dot{x}_i = x_i(r_i - s_i x_i) + x_i \sum_{j \neq i} A_{ij} x_j \quad i = 1, \dots, N \quad (\text{S11})$$

The above equations take into account species-species interactions and self-regulation, this latter effect being described in terms of a logistic growth. Despite its intrinsic simplicity, the model is often invoked to explain how ecological interactions, e.g., mutualism and competition for nutrients, can lead to complex phenomena, as multi-stability or antibiotic mediated catastrophic losses of biodiversity. In [40] the model has been applied to a relatively small (mice gut) microbiota system made up of 11 distinct populations. More precisely, the ten most abundant species have been identified: all together they account for the vast majority ( $\sim 90\%$ ) of the total populations found in the mice gut. The remaining populations are grouped into a unique (non-homogeneous) category referred to as “Other”. The authors of [40] analyzed the experiments reported in [41] and provided a quantitative characterization of the coefficients that enter the definition of the relevant quantities  $\mathbf{r}$ ,  $\mathbf{s}$  and  $\mathbf{A}$ . These latter are reported in table (S12) together with the names of the involved species:

Populations	$\mathbf{r}$	$\mathbf{s}$	$\mathbf{A}$											
<i>Barnesiella</i>	0.368	0.205	0	0.0984	0.167	-0.165	-0.143	0.0199	-0.515	-0.392	0.346	0.00888	-0.269	
undefined genus of Lachnospiraceae	0.310	0.105	0.0621	0	-0.0430	-0.155	-0.187	0.0270	-0.459	-0.414	0.301	0.0221	-0.196	
unclassified Lachnospiraceae	0.356	0.102	0.144	-0.192	0	-0.140	-0.165	0.0136	-0.504	-0.772	0.292	-0.00596	-0.206	
Other	0.540	0.831	0.224	0.138	0.000459	0	-0.224	0.220	-0.205	-1.01	0.666	-0.0390	-0.400	
<i>Blautia</i>	0.709	0.709	-0.180	-0.0513	$-5.03 \times 10^{-5}$	-0.0542	0	0.0162	-0.507	0.554	0.157	0.224	0.10635	
undefined genus of unclassified Mollicutes	0.471	0.423	-0.111	-0.037	-0.0426	0.0410	0.261	0	-0.185	-0.4326	0.165	-0.0610	-0.265	
<i>Akkermansia</i>	0.230	1.21	-0.127	-0.186	-0.122	0.381	0.400	-0.161	0	1.390	-0.379	0.192	-0.0963	
<i>Coprobacillus</i>	0.830	4.35	-0.0712	$6.04 \times 10^{-4}$	0.0803	-0.455	-0.503	0.169	-0.562	0	0.443	-0.223	-0.207	
<i>Clostridium difficile</i>	0.392	0.0558	-0.0375	-0.0333	-0.0499	-0.0904	-0.102	0.0323	-0.182	-0.303	0	0.0144	-0.00767	
<i>Enterococcus</i>	0.291	0.192	-0.0422	-0.0131	0.0240	-0.118	-0.329	0.0207	0.0548	-2.10	0.111	0	0.0238	
undefined genus of Enterobacteriaceae	0.324	0.384	-0.374	0.278	0.249	-0.168	0.0840	0.0337	-0.232	-0.395	0.314	-0.0388	0	

(S12)

The concentration of the species are measured in  $10^{11}$  rRNAcopies/cm<sup>3</sup>, the coefficients  $\mathbf{r}$  have the dimension of the inverse of time (measured in days), and  $\mathbf{s}$  is expressed as the inverse of the product of a time for a concentration.

The second application discussed in the main body of the paper, i.e. that targeted to controlling the microbiota dynamics, assumes the above parameters. Notice that the set of considered species includes the spore-forming pathogen *Clostridium difficile*. To lower its concentration (and so diminish the probability of infection) is one of the goals of the implemented control. As explained in the paper, three different control schemes are considered. In the following we will provide some additional information for each of the analyzed schemes. The physical dimension of the inserted controller  $u$  is again  $10^{11}$  rRNAcopies/cm<sup>3</sup>. The parameters  $\beta$  have dimension of the inverse of a time, while  $\rho$  is a-dimensional.

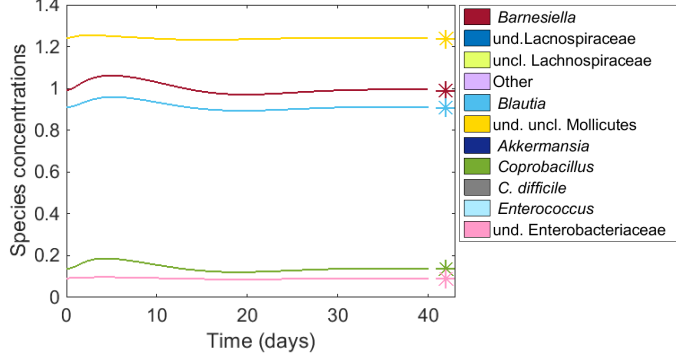


FIG. 3: Numerical integration of the controlled system (see equations (2) in the main body). The equilibrium state stabilized upon injection of the controller (stars) is a slight modified version of the initially unstable fixed point, see table (S13). The system is initialized out of equilibrium and, after a transient, converges to  $\mathbf{x}^*$ .

**Case A: Stabilizing an unstable fixed point by means of an external controller.** The first of the three different control strategies for the microbiota network discussed in the main body involves only 5 out of the 11 populations being examined (more precisely *Barnesiella*, *Blautia*, und. Mollicutes, *Coprobacillus* and und. Enterobacteriaceae). An unstable fixed point of equation (S11), which contains only these latter species, has been calculated and denoted by  $\bar{\mathbf{x}}$  (see table (S13)). The maximum eigenvalue of the Jacobian matrix evaluated at  $\bar{\mathbf{x}}$  is  $(\lambda_{Re})_{max} = 0.0148 > 0$ , thus implying instability. As discussed in the main body of the paper, it is possible to stabilize a different equilibrium, denoted by  $\mathbf{x}^*$  in (S13), which corresponds to a slight modification of the unstable solution  $\bar{\mathbf{x}}$ . Following the procedure detailed in the paper one can readily calculate the parameters  $\alpha$  and  $\beta$  (S13). The values obtained are reported in table (S13). Direct simulations of the controlled system, as displayed in Figure 3, confirms that stability has been indeed achieved.

Populations	$\bar{\mathbf{x}}$	$\mathbf{x}^*$	$\alpha$	$\beta \times 10^4$
<i>Barnesiella</i>	0.9736	0.9917	0.0186	0.1860
<i>Blautia</i>	0.8840	0.9093	0.0089	-0.0638
und. uncl. Mollicutes	1.2361	1.2396	0.0089	-0.0032
<i>Coprobacillus</i>	0.1169	0.1363	0.1005	0.0495
und. Enterobacteriaceae	0.0756	0.0894	0.0175	-2.2067

(S13)

**Case B: Acting with one species of the pool to damp the concentration of the pathogens.** Consider the stable fixed point  $\mathbf{x}^*$  (table (S14)) composed by *Barnesiella*, und. Lachnospiraceae, Other, *Blautia* and *C. difficile*. As explained in the paper, we now insert in the system another population, selected from the extended pool of interacting species. This latter configures as the controller. Vector  $\alpha$  therefore follows from the interaction matrix  $\mathbf{A}$ . One can then calculate the fixed point that can be eventually attained by the system, given the specific selected controller. Retaining only the meaningful cases (fixed points with all positive entries), we obtain three possible solutions:  $\mathbf{x}_L^*$  where the added species is uncl. Lachnospiraceae,  $\mathbf{x}_M^*$  adding as external control the und. uncl. Mollicutes and  $\mathbf{x}_E^*$  adding und. Enterobacteriaceae (see table S14). From inspection of the obtained solutions, one can appreciate the impact of the different controllers employed: in the latter case the concentration of *C. difficile* stays almost constant, in the second example it increases, while in the first case it is reduced by a significant amount.

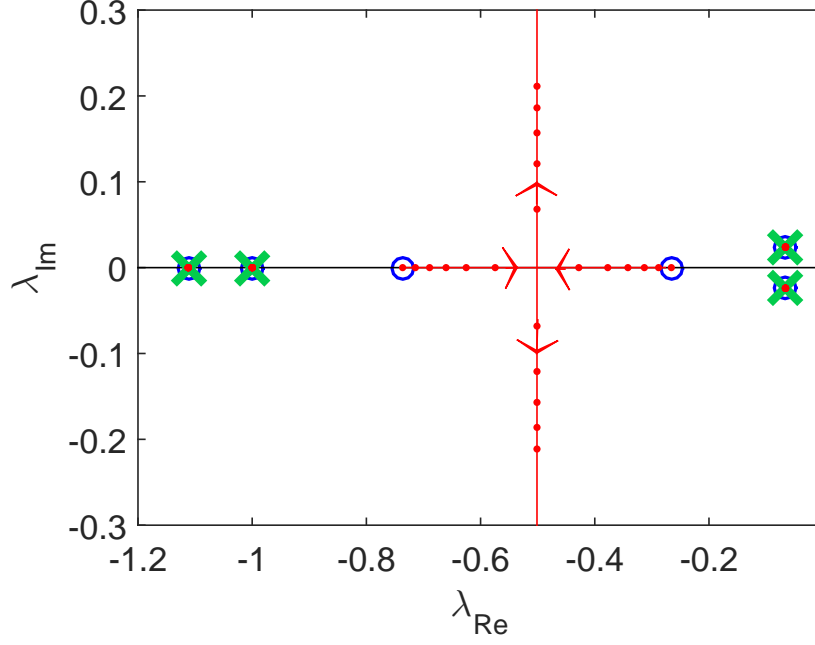


FIG. 4: Root locus diagram relative to the Case B where, starting from the pool of 5 populations the control is performed adding the species of uncl. Lachnospiraceae. Blue circles correspond to the position in the complex plane of the roots of  $\mathcal{D}(\lambda)$  (eigenvalues of  $\mathbf{J}$  when  $\rho = 0$ ) while green crosses indicate the eigenvalues of the Jacobian in the limit  $\rho \rightarrow \infty$ . Paths followed by the eigenvalues when progressively increasing  $\rho$  are shown by the red lines, while the red dots represent the solutions for discrete values of  $\rho$ , scanning the interval from 0 to 0.1.

Populations	$\mathbf{x}^*$	$\mathbf{x}_L^*$	$\mathbf{x}_M^*$	$\mathbf{x}_E^*$
<i>Barnesiella</i>	2.0745	2.5166	2.0480	2.2542
und. Lachnospiraceae	2.3607	1.9876	2.2674	2.7007
Other	1.9608	1.6929	1.9252	2.0842
<i>Blautia</i>	0.2724	0.1422	0.6171	0.1345
<i>C. difficile</i>	0.5402	0.2435	0.6194	0.5259

(S14)

**Case C: Driving to extinction one species, the other being the target of the control.** In this case we aim at enforcing the extinction of one of the species (here *C. difficile*), acting on the other ones. In other words the parameter vector  $\alpha$  (which characterizes the action of the control against the species) is forced to have the component relative to *C. difficile* equal to zero. At the same time, the corresponding entry of the vector  $\mathbf{x}^*$  to be eventually stabilized is set to a negligible value. The concentrations of the species for respectively the initial and the final fixed points are compared in Figure 2 of the main body. The corresponding values of  $\alpha$  are also plotted. As a supplementary material, we here report in Figure 5 the root locus diagram obtained for the case at hand.

### Exploiting a transient control to drive the system towards an existing stable fixed point

As anticipated in the main body of the paper, the control can be effectively employed to steer the system towards a stable fixed point of the unperturbed dynamics, starting from out-of-equilibrium initial conditions. In this case, we set  $\dot{u} = -\gamma u - \sum_j \beta_j (x_j - x_j^*)$ . Here,  $x_j^*$  is an equilibrium solution of the uncontrolled dynamics, which proves linearly stable to external perturbations. The parameter  $\gamma$  can be tuned as desired so as to help the convergence towards  $x_j^*$  without falling in the basin of attraction of other existing fixed points. As a proof of principle of the method, we

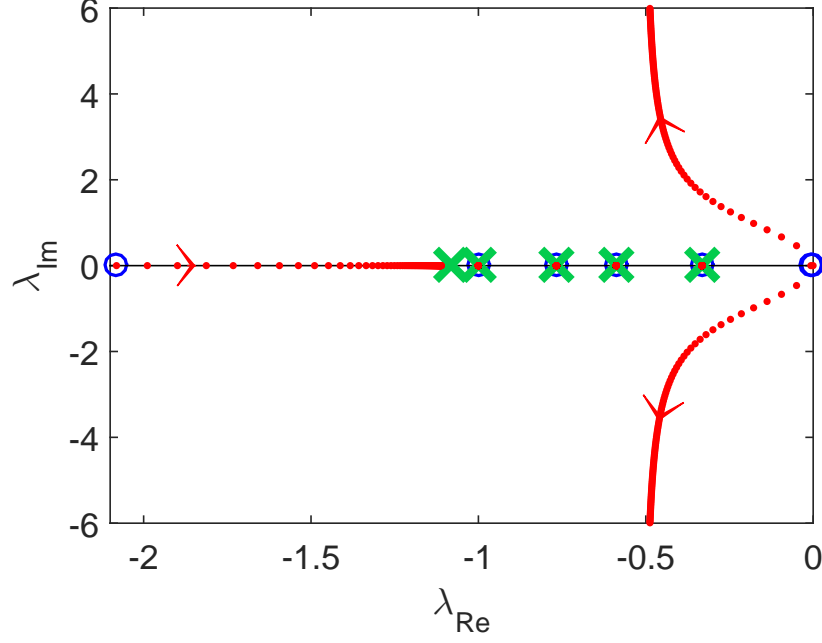


FIG. 5: Root locus diagram relative to the Case C, where an indirect control is applied on the reduced system of 6 species. For a description of the symbols, refer to the caption of Figure 4.

choose a stable fixed point of the global microbiota ecosystem, i.e. including the complete pool of 11 populations. This is characterized by  $x_1^* = 9.299$ ,  $x_3^* = 12.3085$ ,  $x_4^* = 3.1627$  and  $x_j^* = 0$  for  $j \neq 1, 3, 4$ . The largest real part of the eigenvalues of the associated  $11 \times 11$  Jacobian matrix turns out to be  $(\lambda_{Re}) = -0.1306 < 0$ , thus implying stability of the aforementioned equilibrium. Imagine to initialize the system out of equilibrium with all species, including the pathogen *C. difficile*, being assigned a random concentration  $x_j(0) \neq 0$ . The system is let evolve for a while and then, at time  $t^*$ , the control is injected. Here,  $\alpha$  and  $\beta$  are assigned as random, uniformly distributed over  $[0,1]$ , parameters. As clearly displayed in Figure 6, the system is steadily moved towards the equilibrium  $x_j^*$  (stars), while the control converges to zero after an abrupt jump. In other words, after a transient whose duration depends on the chosen parameters, the system achieves its asymptotic (pathogen free) equilibrium and the control can be safely disconnected.

### On the controllability condition

Observe that the control procedure here discussed requires computing the vector  $\beta$ . Determining this latter implies inverting a matrix, an operation that imposes a mathematical constraint that we shall hereafter analyze more in depth. The matrix to be inverted, as defined in the main text, reads:

$$H_{nm} = \sum_{k=0}^{N-n} c_{k+n+1} (G^k \mathbf{q})_m \quad (\text{S15})$$

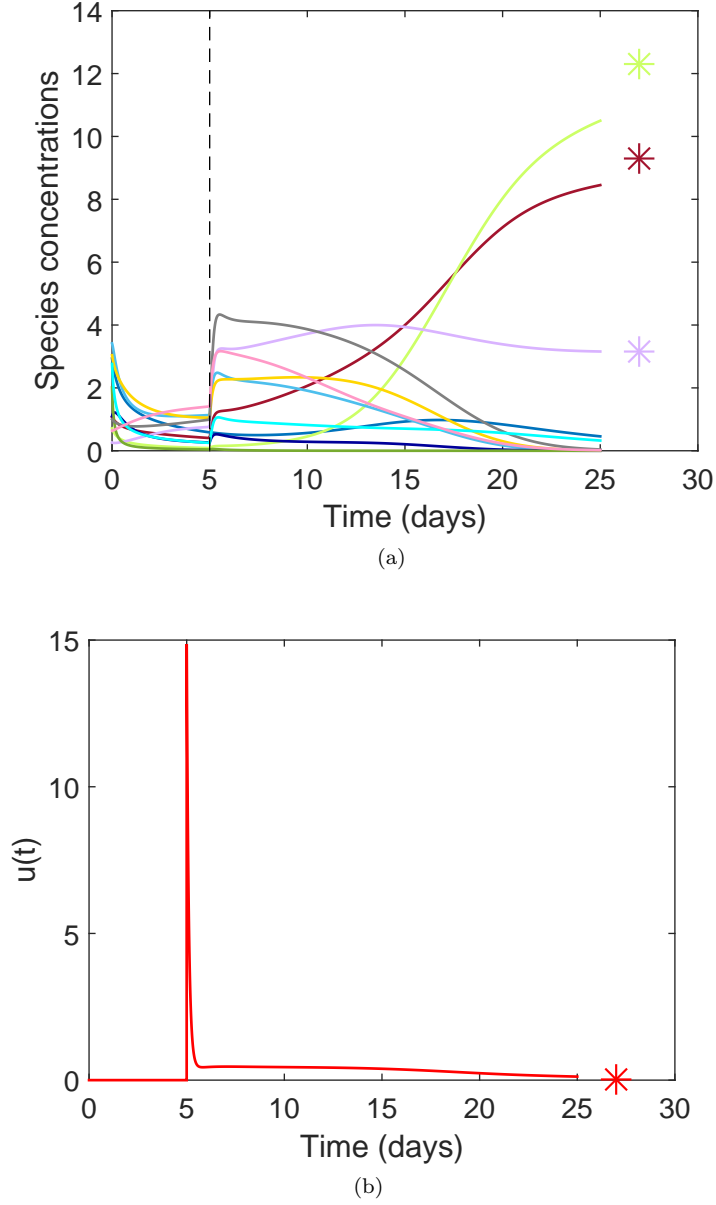


FIG. 6: Driving the system towards a stable fixed point of the unperturbed dynamics. The system is initiated out-of-equilibrium: the concentration of all species, including the pathogen *C. difficile*, is set to values different from zero. At  $t^* = 5$  days the control  $u$  is injected. After a transient the system converges to its stable equilibrium characterized by  $x_1^* = 9.299$ ,  $x_3^* = 12.3085$ ,  $x_4^* = 3.1627$  and  $x_j^* = 0$  for  $j \neq 1, 3, 4$ , while the control  $u$  is turned to zero. Here  $\gamma = 10$  and  $u(0) = 15$ .

where  $c$  stands for the coefficients<sup>4</sup> of the characteristic polynomial of  $\mathbf{G}$ , namely  $\det(\mathbf{G} - \lambda \mathbf{I}) = \sum_{l=0}^N c_{l+1} \lambda^l$ . The invertibility is ensured if  $\det(\mathbf{H}) \neq 0$ , in formulae:

$$\begin{aligned} \det(\mathbf{H}) &= \sum_{i_1=1}^N \cdots \sum_{i_N=1}^N \epsilon_{i_1, \dots, i_N} H_{1i_1} H_{2i_2} \cdots H_{N-1i_{N-1}} H_{Ni_N} = \\ &= \sum_{i_1=1}^N \cdots \sum_{i_N=1}^N \epsilon_{i_1, \dots, i_N} \left[ \sum_{k_1=0}^{N-1} c_{k_1+2} (\mathbf{G}^{k_1} \mathbf{q})_{i_1} \right] \left[ \sum_{k_2=0}^{N-2} c_{k_2+3} (\mathbf{G}^{k_2} \mathbf{q})_{i_2} \right] \cdots \\ &\cdots \left[ \sum_{k_{N-1}=0}^1 c_{k_{N-1}+N} (\mathbf{G}^{k_{N-1}} \mathbf{q})_{i_{N-1}} \right] [c_{N+1} q_{i_N}] \neq 0 \end{aligned} \quad (\text{S17})$$

where  $\epsilon_{i_1, \dots, i_N}$  is the Levi-Civita tensor. This complicated expression can be heavily simplified. The Levi-Civita symbol is in fact totally antisymmetric in the permutation of its indices. As a consequence, all the terms multiplied by  $\epsilon_{i_1, \dots, i_N}$  which are symmetric in the permutations, cancel out. It follows that the terms containing the product of two or more factors  $(\mathbf{G}^k \mathbf{q})$  with the same power  $k$  are identically equal to zero. The only terms which survive are those obtained by just retaining the largest possible value of  $k$  in each summation ( $k_1 = N-1$ ,  $k_2 = N-2, \dots$ ,  $k_{N-1} = 1$ ). In formulae:

$$\begin{aligned} \det(\mathbf{H}) &= \sum_{i_1=1}^N \cdots \sum_{i_N=1}^N \epsilon_{i_1, \dots, i_N} c_{N+1} (\mathbf{G}^{N-1} \mathbf{q})_{i_1} c_{N+1} (\mathbf{G}^{N-2} \mathbf{q})_{i_2} \cdots c_{1+N} (\mathbf{G} \mathbf{q})_{i_{N-1}} c_{N+1} q_{i_N} = \\ &= (-1)^{N^2} \sum_{i_1=1}^N \cdots \sum_{i_N=1}^N \epsilon_{i_1, \dots, i_N} (\mathbf{G}^{N-1} \mathbf{q})_{i_1} (\mathbf{G}^{N-2} \mathbf{q})_{i_2} \cdots (\mathbf{G} \mathbf{q})_{i_{N-1}} q_{i_N}. \end{aligned} \quad (\text{S18})$$

where use has been made of the fact that  $c_{N+1} = (-1)^N$  (see equation (S16) in the footnote 4).

Drawing on this preliminary observations it is possible to re-interpret the above controllability constraint, making contact with standard control theory. The controllability condition amounts to require that the matrix

$$\mathbf{C} \equiv [\mathbf{q}, \mathbf{G}\mathbf{q}, \dots, \mathbf{G}^{N-1}\mathbf{q}] = \begin{pmatrix} q_1 & (\mathbf{G}\mathbf{q})_1 & (\mathbf{G}^2\mathbf{q})_1 & \cdots & (\mathbf{G}^{N-1}\mathbf{q})_1 \\ q_2 & (\mathbf{G}\mathbf{q})_2 & (\mathbf{G}^2\mathbf{q})_2 & \cdots & (\mathbf{G}^{N-1}\mathbf{q})_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_N & (\mathbf{G}\mathbf{q})_N & (\mathbf{G}^2\mathbf{q})_N & \cdots & (\mathbf{G}^{N-1}\mathbf{q})_N \end{pmatrix} \quad (\text{S19})$$

has maximum rank. Here, with the notation  $(\mathbf{G}^l \mathbf{q})_k$  we identify the  $k$ -th entry of the vector obtained from the product of the matrix  $\mathbf{G}$  to the power of  $l$  with vector  $\mathbf{q}$ . In system theory the matrix  $\mathbf{C}$  is called the *controllability* matrix of the pair  $(\mathbf{G}, \mathbf{q})$ . Since  $\mathbf{C}$  is a square matrix, the maximum rank condition is equivalent to require  $\det(\mathbf{C}) \neq 0$ , in

---

4

$$\begin{aligned} c_1 &= \det(\mathbf{G}) \\ c_k &= \frac{(-1)^{k-1}}{(k-1)!} \sum_{\substack{j_1 \dots j_{k-1}=1 \\ j_1 \neq \dots \neq j_{k-1}}}^N \det \mathbf{G}_{(j_1 j_1)(j_2 j_2) \dots (j_{k-1} j_{k-1})} \text{ for } k = 2, \dots, N-1 \\ c_N &= (-1)^{N-1} \text{Tr}(\mathbf{G}) \\ c_{N+1} &= (-1)^N \end{aligned} \quad (\text{S16})$$

where with  $\mathbf{G}_{(j_1 j_1)(j_2 j_2) \dots (j_{k-1} j_{k-1})}$  we identify the minor obtained from matrix  $\mathbf{G}$  by removing the  $j_1$ -th,  $j_2$ -th, ...,  $j_{k-1}$ -th rows and columns.

formulae:

$$\begin{aligned}
\det(\mathbf{C}) &= \sum_{\sigma_1=1}^N \cdots \sum_{\sigma_N=1}^N \epsilon_{\sigma_1, \dots, \sigma_N} C_{\sigma_1 1} C_{\sigma_2 2} \cdots C_{\sigma_{N-1} N-1} C_{\sigma_N N} = \\
&= \sum_{\sigma_1=1}^N \cdots \sum_{\sigma_N=1}^N \epsilon_{\sigma_1, \dots, \sigma_N} (\mathbf{G}^0 \mathbf{q})_{\sigma_1} (\mathbf{G}^1 \mathbf{q})_{\sigma_2} \cdots (\mathbf{G}^{N-2} \mathbf{q})_{\sigma_{N-1}} (\mathbf{G}^{N-1} \mathbf{q})_{\sigma_N} = \\
&= (-1)^N \sum_{\sigma_1=1}^N \cdots \sum_{\sigma_N=1}^N \epsilon_{\sigma_1, \dots, \sigma_N} (\mathbf{G}^{N-1} \mathbf{q})_{\sigma_1} (\mathbf{G}^{N-2} \mathbf{q})_{\sigma_2} \cdots (\mathbf{G} \mathbf{q})_{\sigma_{N-1}} (\mathbf{q})_{\sigma_N} \neq 0
\end{aligned} \tag{S20}$$

where use has been made of the definition of  $\mathbf{C}$  (S19), namely  $C_{ij} = (\mathbf{G}^{j-1} \mathbf{q})_i$ .

Expressions (S18) and (S20) are identical (except for the sign) and consequently the two conditions,  $\det(\mathbf{H}) \neq 0$  e  $\det(\mathbf{C}) \neq 0$ , prove equivalent. Stated differently, the condition of invertibility of matrix  $\mathbf{H}$ , obtained as a self-consistent constraint for the introduced control scheme, coincides with the standard controllability condition, as known in control theory.

As a final remark we recall (see main body) that our goal is not to arbitrarily assign the polynomial  $\mathcal{N}(\lambda)$  but rather to locate its roots  $z_k$  within the open left-hand plane. In this respect, the necessary and sufficient system-theoretic condition is the so-called *stabilizability* of the pair  $(\mathbf{G}, \mathbf{q})$ , which results in the Popov-Belevitch-Hautus rank condition [33, 44].

- 
- [1] J. Murray, “Mathematical biology ii: Spatial models and biochemical applications, volume ii,” 2003.
  - [2] M. C. Cross and P. C. Hohenberg, “Pattern formation outside of equilibrium,” *Reviews of modern physics*, vol. 65, no. 3, p. 851, 1993.
  - [3] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: a universal concept in nonlinear sciences*, vol. 12. Cambridge university press, 2003.
  - [4] S. Suweis, F. Simini, J. R. Banavar, and A. Maritan, “Emergence of structural and dynamical properties of ecological mutualistic networks,” *Nature*, vol. 500, no. 7463, pp. 449–452, 2013.
  - [5] G. Caldarelli and A. Chessa, *Data Science and Complex Networks: Real Case Studies with Python*. Oxford University Press, 2016.
  - [6] K. Z. Coyte, J. Schluter, and K. R. Foster, “The ecology of the microbiome: Networks, competition, and stability,” *Science*, vol. 350, no. 6261, pp. 663–666, 2015.
  - [7] E. Kandel, J. Schwartz, and T. Jessell, *Principles of Neural Science*. McGraw-Hill, fourth ed., 2000.
  - [8] M. Asllani, J. D. Challenger, F. S. Pavone, L. Sacconi, and D. Fanelli, “The theory of pattern formation on directed networks,” *Nature communications*, vol. 5, 2014.
  - [9] H. Lodish, A. Berk, S. L. Z., P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology*. New York: W. H. Freeman, fourth ed., 2000.
  - [10] M. Rohden, A. Sorge, M. Timme, and D. Witthaut, “Self-organized synchronization in decentralized power grids,” *Phys. Rev. Lett.*, vol. 109, p. 064101, 2012.
  - [11] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, “The structure and dynamics of multilayer networks,” *Physics Reports*, vol. 544, no. 1, pp. 1–122, 2014.
  - [12] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports*, vol. 424, pp. 175–308, 2006.
  - [13] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks*. Cambridge: Cambridge University Press, 2008.
  - [14] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, “Synchronization in complex networks,” *Phys. Rep.*, vol. 469, pp. 93–153, 2008.
  - [15] J. Gao, B. Barzel, and A.-L. Barabási, “Universal resilience patterns in complex networks,” *Nature*, vol. 530, no. 7590, pp. 307–312, 2016.
  - [16] J. Grilli, T. Rogers, and S. Allesina, “Modularity and stability in ecological communities,” *Nature communications*, vol. 7, 2016.
  - [17] E. Thébault and C. Fontaine, “Stability of ecological communities and the architecture of mutualistic and trophic networks,” *Science*, vol. 329, no. 5993, pp. 853–856, 2010.
  - [18] F. Dorfler and F. Bullo, “Synchronization and transient stability in power networks and nonuniform kuramoto oscillators,” *SIAM Journal on Control and Optimization*, vol. 50, no. 3, pp. 1616–1642, 2012.
  - [19] G. Cencetti, F. Bagnoli, G. Battistelli, L. Chisci, F. Di Patti, and D. Fanelli, “Topological stabilization for synchronized dynamics on networks,” *The European Physical Journal B*, vol. 90, no. 1, p. 9, 2017.

- [20] R. E. Kalman, "Mathematical description of linear dynamical systems," *Journal of the Society for Industrial and Applied Mathematics Series A Control*, vol. 1, no. 2, pp. 152–192, 1963.
- [21] D. Luenberger, "Introduction to dynamic systems: theory, models, and applications," 1979.
- [22] J.-J. E. Slotine, W. Li, *et al.*, *Applied nonlinear control*, vol. 199. Prentice-Hall Englewood Cliffs, NJ, 1991.
- [23] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.
- [24] V. Nicosia, R. Criado, M. Romance, G. Russo, and V. Latora, "Controlling centrality in complex networks," *arXiv preprint arXiv:1109.4521*, 2011.
- [25] C. Tu, J. Grilli, and S. Suweis, "Universality of resilience patterns in generalized lotka volterra dynamics and beyond," *arXiv preprint arXiv:1606.09630*, 2016.
- [26] R. M. May, "Will a large complex system be stable?," *Nature*, vol. 238, pp. 413–414, 1972.
- [27] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, no. 6786, pp. 590–593, 2000.
- [28] K. A. Johnson and R. S. Goody, "The original michaelis constant: translation of the 1913 michaelis–menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011.
- [29] M. Cavallaro, D. Asprone, V. Latora, G. Manfredi, and V. Nicosia, "Assessment of urban ecosystem resilience through hybrid socialphysical complex networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 29, no. 8, pp. 608–625, 2014.
- [30] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, and V. Latora, "Remote synchronization reveals network symmetries and functional modules," *Phys. Rev. Lett.*, vol. 110, p. 174102, Apr 2013.
- [31] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *Science*, vol. 345, no. 6198, pp. 795–799, 2014.
- [32] W. R. Evans, "Graphical analysis of control systems," *Transactions of the American Institute of Electrical Engineers*, vol. 67, no. 1, pp. 547–551, 1948.
- [33] T. Kailah, *Linear Systems*. Prentice-Hall Englewood Cliffs, NJ, 1980.
- [34] R. M. Albert J., "Dynamic modeling of gene expression in prokaryotes: application to glucose-lactose diauxie in escherichia coli," *Systems and Synthetic Biology*, vol. 5, no. 1, pp. 33–43, 2011.
- [35] J. Hasty, D. McMillen, F. Isaacs, and J. J. Collins, "Computational studies of gene regulatory networks: in numero molecular biology," *Nature Reviews Genetics*, vol. 2, no. 4, pp. 268–279, 2001.
- [36] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.
- [37] F. J. Isaacs, J. Hasty, C. R. Cantor, and J. J. Collins, "Prediction and measurement of an autoregulatory genetic module," *Proceedings of the National Academy of Sciences*, vol. 100, no. 13, pp. 7714–7719, 2003.
- [38] E. Signaroldi, P. Laise, S. Cristofanon, A. Brancaccio, E. Reisoli, S. Atashpaz, M. R. Terreni, C. Doglioni, G. Pruneri, P. Malatesta, *et al.*, "Polycomb dysregulation in gliomagenesis targets a zfp423-dependent differentiation network," *Nature communications*, vol. 7, 2016.
- [39] T.-C. D. Shen, L. Albenberg, K. Bittinger, C. Chehoud, Y.-Y. Chen, C. A. Judge, L. Chau, J. Ni, M. Sheng, A. Lin, B. J. Wilkins, E. L. Buza, J. D. Lewis, Y. Daikhin, I. Nissim, M. Yudkoff, F. D. Bushman, and G. D. Wu, "Engineering the gut microbiota to treat hyperammonemia," *The Journal of Clinical Investigation*, vol. 125, pp. 2841–2850, 7 2015.
- [40] R. R. Stein, V. Bucci, N. C. Toussaint, C. G. Buffie, G. Rätsch, E. G. Pamer, C. Sander, and J. B. Xavier, "Ecological modeling from time-series inference: Insight into dynamics and stability of intestinal microbiota," *PLOS Computational Biology*, vol. 9, pp. 1–11, 12 2013.
- [41] C. G. Buffie, I. Jarchum, M. Equinda, L. Lipuma, A. Gobourne, A. Viale, C. Ubeda, J. Xavier, and E. G. Pamer, "Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to clostridium difficile-induced colitis," *Infection and immunity*, vol. 80, no. 1, pp. 62–73, 2012.
- [42] S. N. Steinway, M. B. Biggs, T. P. Loughran Jr, J. A. Papin, and R. Albert, "Inference of network dynamics and metabolic interactions in the gut microbiome," *PLoS Comput Biol*, vol. 11, no. 6, p. e1004338, 2015.
- [43] S. Freilich, R. Zarecki, O. Eilam, E. S. Segal, C. S. Henry, M. Kupiec, U. Gophna, R. Sharan, and E. Rupp, "Competitive and cooperative metabolic interactions in bacterial communities," *Nature communications*, vol. 2, p. 589, 2011.
- [44] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang, and Y.-C. Lai, "Exact controllability of complex networks," *Nature communications*, vol. 4, 2013.